

Elsevier L^AT_EX template[☆]

Elsevier¹

Radarweg 29, Amsterdam

*Elsevier Inc^{a, b}, Global Customer Service^{b, *}*

^a1600 John F Kennedy Boulevard, Philadelphia

^b360 Park Avenue South, New York

Abstract

This template helps you to create a properly formatted L^AT_EX manuscript.

Keywords:

1. Main

Suppose $(X_1^T, Y_1), \dots, (X_n^T, Y_n)$ are i.i.d. from $N_{p+1}(\mu, \Sigma)$, where $X_i \in \mathbb{R}^p$ and $Y_i \in \mathbb{R}$. Denote $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_n)^T$.

Write $Y = \beta_0 \mathbf{1}_n + X^T \beta + \epsilon$, where $\mathbf{1}_n$ is n dimensional vector with all elements equal to 1. ϵ has distribution $N(0, \sigma^2 I_n)$.

The problem is to test hypotheses $H : \beta = 0$.

Let $Q_n = WW^T$ be the rank decomposition of Q_n , where W_n is a $n \times n - 1$ matrix with $W^T W = I_{n-1}$. The new test statistic is

$$T = \frac{y^T Q_n y}{y^T W (W^T X^T X W)^{-1} W^T y}$$

or equivalently

$$\frac{y^T Q_n y}{y^T Q_n (X^T X)^{-1} Q_n y - (y^T Q_n (X^T X)^{-1} \mathbf{1}_n)^2 / (\mathbf{1}_n^T (X^T X)^{-1} \mathbf{1}_n)}$$

[☆]Fully documented templates are available in the elsarticle package on CTAN.

^{*}Corresponding author

Email address: support@elsevier.com (Global Customer Service)

URL: www.elsevier.com (Elsevier Inc)

¹Since 1880.

Let $\tilde{y} = W^T y$, $\tilde{X} = XW$, $\tilde{\epsilon} = W^T \epsilon$. Then

$$\tilde{y} = \tilde{X}^T \beta + \tilde{\epsilon}$$

and

$$T = \frac{\tilde{y}^T \tilde{y}}{\tilde{y}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{y}}$$

Next we derive another form of T . We follow the similar technique of Hotelling's T^2 .

Let R be an $(n-1) \times (n-1)$ orthogonal matrix satisfies

$$R\tilde{y} = \begin{pmatrix} \|\tilde{y}\| \\ 0 \\ \dots \\ 0 \end{pmatrix}.$$

We can write

$$T = \frac{\|\tilde{y}\|^2}{\tilde{y}^T R^T (R \tilde{X}^T \tilde{X} R^T)^{-1} R \tilde{y}} \quad (1)$$

Denote by $B = R \tilde{X}^T \tilde{X} R^T$, then

$$T = \frac{1}{(B^{-1})_{11}}.$$

Let

$$B = \begin{pmatrix} b_{11} & b_{(1)}^T \\ b_{(1)} & B_{22} \end{pmatrix},$$

and apply the matrix inverse formula, we have $(B^{-1})_{11} = 1/(b_{11} - b_{(1)}^T B_{22}^{-1} b_{(1)})$.

Hence

$$T = b_{11} - b_{(1)}^T B_{22}^{-1} b_{(1)}.$$

2. Asymptotic distribution

Note that conditioning on \tilde{y} , R is a constant orthogonal matrix. And \tilde{y} is independent of \tilde{X} under null hypotheses. So $B|\tilde{y}$ has the same distribution with

$\tilde{X}^T \tilde{X}$ under null hypotheses. Hence B is independent of \tilde{y} and can be written as

$$B = \sum_{i=1}^p \lambda_i z_i z_i^T \quad (2)$$

where z_i 's are i.i.d. $n-1$ dimensional random vectors distributed as $N(0, I_{n-1})$, $\lambda_1 \geq \lambda_2 \dots \geq \lambda_p > 0$ are eigenvalues of Σ_X . Denote by $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$, $Z = (Z_1, \dots, Z_p)$. Let $Z_{(1)}$ and $Z_{(2)}$ be the first 1 row and last $n-2$ rows of Z , that is

$$Z = \begin{pmatrix} Z_{(1)} \\ Z_{(2)} \end{pmatrix}.$$

Then

$$\begin{aligned} B &= Z \Lambda Z^T \\ &= \begin{pmatrix} Z_{(1)} \Lambda Z_{(1)}^T & Z_{(1)} \Lambda Z_{(2)}^T \\ Z_{(2)} \Lambda Z_{(1)}^T & Z_{(2)} \Lambda Z_{(2)}^T \end{pmatrix}. \end{aligned} \quad (3)$$

Hence

$$\begin{aligned} T &= Z_{(1)} \Lambda Z_{(1)}^T - Z_{(1)} \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda Z_{(1)}^T \\ &= Z_{(1)} (\Lambda - \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) Z_{(1)}^T. \end{aligned} \quad (4)$$

But

$$\begin{aligned} \text{rank}(\Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) &= \text{rank}(\Lambda^{\frac{1}{2}} Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda^{\frac{1}{2}}) \\ &= \text{rank}(I_{n-2}) = n-2, \end{aligned} \quad (5)$$

and

$$\begin{aligned} \text{rank}(\Lambda - \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) &= \text{rank}(I_p - \Lambda^{\frac{1}{2}} Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda^{\frac{1}{2}}) \\ &= p - n + 2. \end{aligned} \quad (6)$$

Hence

$$T \sim \sum_{i=1}^{p-n+2} \lambda_i (\Lambda - \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) \chi_1^2$$

By Weyl's inequality, we have for $1 \leq i \leq p-n+2$

$$\lambda_i (\Lambda - \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) \leq \lambda_i (\Lambda), \quad (7)$$

and

$$\begin{aligned}
& \lambda_i (\Lambda - \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) \\
& \geq \lambda_{i+n-2} (\Lambda) + \lambda_{p-n+2} (-\Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda) \\
& = \lambda_{i+n-2}.
\end{aligned} \tag{8}$$

Hence

$$\sum_{i=n-1}^p \lambda_i \chi_1^2 \leq T \leq \sum_{i=1}^{p-n+2} \lambda_i \chi_1^2$$

Note that under condition $\text{tr} \Sigma^4 / (\text{tr} \Sigma^2)^2 \rightarrow 0$, we have by Liapounoff central limit theorem that

$$\frac{\sum_{i=1}^p \lambda_i \chi_1^2 - \text{tr} \Sigma_X}{\sqrt{\text{tr}(\Sigma_X^2)}} \xrightarrow{\mathcal{L}} N(0, 1).$$

And

$$\frac{T - \text{tr} \Sigma_X}{\sqrt{\text{tr}(\Sigma_X^2)}} - \frac{\sum_{i=1}^p \lambda_i \chi_1^2 - \text{tr} \Sigma_X}{\sqrt{\text{tr}(\Sigma_X^2)}} = \frac{T - \sum_{i=1}^p \lambda_i \chi_1^2}{\sqrt{\text{tr}(\Sigma_X^2)}}, \tag{9}$$

To prove (9) $\xrightarrow{P} 0$, we only need to prove

$$\mathbb{E} \left(\frac{\sum_{i=1}^{n-2} \lambda_i \chi_1^2}{\sqrt{\text{tr}(\Sigma_X^2)}} \right) \rightarrow 0,$$

that is

$$\mathbb{E} \left(\frac{\sum_{i=1}^{n-2} \lambda_i}{\sqrt{\text{tr}(\Sigma_X^2)}} \right) \rightarrow 0. \tag{10}$$

If λ_i 's are bounded below and above, then (10) is equivalent to

$$n / \sqrt{p} \rightarrow 0, \tag{11}$$

or $p/n^2 \rightarrow \infty$. We thus obtain the following theorem.

Theorem 1. *Suppose*

$$\mathbb{E} \left(\frac{\sum_{i=1}^{n-2} \lambda_i}{\sqrt{\text{tr}(\Sigma_X^2)}} \right) \rightarrow 0,$$

and

$$\frac{\text{tr} \Sigma^4}{(\text{tr} \Sigma^2)^2} \rightarrow 0.$$

Then under null hypotheses, we have

$$\frac{T - \text{tr} \Sigma_X}{\sqrt{\text{tr}(\Sigma_X^2)}} \xrightarrow{\mathcal{L}} N(0, 1).$$

3. Full Asymptotic Results

$$\Sigma_X = P\Lambda P^T$$

Non-spike: there's no principal component ($r = 0$). That is, $\lambda_1 = \dots = \lambda_p$.

Spike: there's r principal components. That is, $\lambda_1 \geq \lambda_2 \geq \dots \lambda_r \geq \lambda_{r+1} = \dots = \lambda_p$. Denote by P_1 the first r column of P and P_2 the last $p - r$ column of P .

$$\begin{aligned} Y &= \beta_0 \mathbf{1}_n + X^T \beta + \epsilon \\ &= \beta_0 \mathbf{1}_n + X^T P_1 P_1^T \beta + X^T P_2 P_2^T \beta + \epsilon \end{aligned} \tag{12}$$

In either case, let λ be $\lambda = \lambda_{r+1} = \dots = \lambda_p$.

PCR try to do regression between Y and (estimated) $X^T P_1$. If P_1 is observed, then the problem is reduced to testing an ordinary regression model. However, it's not the case.

Simply estimating P_1 and invoke classical testing procedure may not be a good idea since the estimation may not be consistent in high dimension. In fact, there may be even no principal component!

In this paper, testing PCR means testing:

H_0 : There's no principal component or there's r principal components but $P_1^T \beta = 0$.

H_1 : There's r principal components and $P_1^T \beta \neq 0$.

Next we consider:

1. There's no PC.
2. There's r principal components but $P_1 \beta = 0$.

3.1. circumstance 1

$$T = \frac{\beta^T \tilde{X} \tilde{X}^T \beta + 2\beta^T \tilde{X} \tilde{\epsilon} + \tilde{\epsilon}^T \tilde{\epsilon}}{\beta^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \beta + 2\beta^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon} + \tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \tag{13}$$

Independent of data, generate a random p dimensional orthonormal matrix O with Haar invariant distribution. And

$$T = \frac{(O\beta)^T O\tilde{X}(O\tilde{X})^T O\beta + 2(O\beta)^T \tilde{X}\tilde{\epsilon} + \tilde{\epsilon}^T \tilde{\epsilon}}{(O\beta)^T O\tilde{X}((O\tilde{X})^T O\tilde{X})^{-1} (O\tilde{X})^T \beta + 2(O\beta)^T O\tilde{X}((O\tilde{X})^T O\tilde{X})^{-1} \tilde{\epsilon} + \tilde{\epsilon}^T ((O\tilde{X})^T O\tilde{X})^{-1} \tilde{\epsilon}} \quad (14)$$

Note that conditioning on O , $O\tilde{X}$ is a random matrix with each entry independently distributed as $N(0, \lambda)$. Hence O is independent of $O\tilde{X}$. Observe also that $O\beta/\|\beta\|$ is uniformly distributed on the unit ball. We can without loss of generality assuming that $\beta/\|\beta\|$ is uniformly distributed on the unit ball. Independent of data, generate $R > 0$ with R^2 distributed as χ_p^2 . Then $\xi = R\beta/\|\beta\|$ distributed as $N_p(0, I_p)$. Note that conditioning on \tilde{X} , $\eta = (\tilde{X}^T \tilde{X})^{-1/2} \tilde{X}^T \xi$ is distributed as $N_{n-1}(0, I_{n-1})$. Hence η is independent of \tilde{X} .

Then

$$\begin{aligned} T &= \frac{(\|\beta\|/R)^2 \xi^T \tilde{X} \tilde{X}^T \xi + 2(\|\beta\|/R) \xi^T \tilde{X} \tilde{\epsilon} + \tilde{\epsilon}^T \tilde{\epsilon}}{(\|\beta\|/R)^2 \xi^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \xi + 2(\|\beta\|/R) \xi^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon} + \tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \\ &= \frac{(\|\beta\|/R)^2 \eta^T \tilde{X}^T \tilde{X} \eta + 2(\|\beta\|/R) \eta^T (\tilde{X}^T \tilde{X})^{1/2} \tilde{\epsilon} + \tilde{\epsilon}^T \tilde{\epsilon}}{(\|\beta\|/R)^2 \eta^T \eta + 2(\|\beta\|/R) \eta^T (\tilde{X}^T \tilde{X})^{-1/2} \tilde{\epsilon} + \tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \\ &= \frac{A_1 + A_2 + A_3}{B_1 + B_2 + B_3} \end{aligned} \quad (15)$$

We consider

$$\frac{\frac{\|\beta\|^2}{R^2} \eta^T \eta}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} = \frac{(n-1)\|\beta\|^2}{p} \frac{\frac{p}{R^2} \frac{\eta^T \eta}{n-1}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \quad (16)$$

Note that

$$\begin{aligned} \tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon} &\sim \frac{\tilde{\epsilon}^T \tilde{\epsilon}}{(\tilde{X}^T \tilde{X})_{11 \cdot 2}} \\ &\sim \frac{\sigma^2}{\lambda} \frac{\chi_{n-1}^2}{\chi_{p-n+2}^2} \end{aligned} \quad (17)$$

Hence

$$\begin{aligned} \frac{\frac{\|\beta\|^2}{R^2} \eta^T \eta}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} &= \frac{(n-1)(p-n+2)}{p(n-1)} \frac{\lambda}{\sigma^2} \|\beta\|^2 (1 + o_P(1)) \\ &= \frac{\lambda}{\sigma^2} \|\beta\|^2 (1 + o_P(1)) \end{aligned} \quad (18)$$

On the other hand,

$$\frac{\frac{\|\beta\|^2}{R^2} \eta^T \tilde{X}^T \tilde{X} \eta}{\tilde{\epsilon}^T \tilde{\epsilon}} = \frac{\|\beta\|^2}{\sigma^2 p(n-1)} \frac{\frac{p}{R^2} \eta^T \tilde{X}^T \tilde{X} \eta}{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\sigma^2(n-1)}} \quad (19)$$

Note that

$$\begin{aligned} \eta^T \tilde{X}^T \tilde{X} \eta &\sim \eta^T \eta (\tilde{X}^T \tilde{X})_{11} \\ &\sim \lambda \chi_{n-1}^2 \chi_p^2 \end{aligned} \quad (20)$$

Hence

$$\begin{aligned} \frac{\frac{\|\beta\|^2}{R^2} \eta^T \tilde{X}^T \tilde{X} \eta}{\tilde{\epsilon}^T \tilde{\epsilon}} &= \frac{\|\beta\|^2}{\sigma^2 p(n-1)} \lambda(n-1)p(1 + o_P(1)) \\ &= \frac{\lambda}{\sigma^2} \|\beta\|^2 (1 + o_P(1)) \end{aligned} \quad (21)$$

We can deduce that: If $\|\beta\|^2 \rightarrow 0$, then

$$T = \frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} (1 + o_P(1)) \quad (22)$$

If $\|\beta\|^2 \rightarrow \infty$, then

$$T = \frac{\eta^T \tilde{X}^T \tilde{X} \eta}{\eta^T \eta} (1 + o_P(1)) \quad (23)$$

Note that

$$\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \sim (\tilde{X}^T \tilde{X})_{11,2} \sim \lambda \chi_{p-n+2}^2 \quad (24)$$

Hence by CLT,

$$\frac{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} - \lambda(p-n+2)}{\lambda \sqrt{2(p-n+2)}} \xrightarrow{\mathcal{L}} N(0, 1) \quad (25)$$

$$\begin{aligned}
\frac{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} - \lambda(p-n+2)}{\lambda \sqrt{2(p-n+2)}} &= \frac{\sqrt{p}}{\sqrt{p-n+2}} \frac{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} - \lambda(p-n+2)}{\lambda \sqrt{2p}} \\
&= \frac{\sqrt{p}}{\sqrt{p-n+2}} \left(\frac{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} - \lambda p}{\lambda \sqrt{2p}} + \frac{(n-2)}{\sqrt{2p}} \right)
\end{aligned} \tag{26}$$

By slusky Theorem,

$$\frac{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} - \lambda p}{\lambda \sqrt{2p}} \xrightarrow{\mathcal{L}} N(0, 1) \tag{27}$$

On the other hand

$$\frac{\eta^T \tilde{X}^T \tilde{X} \eta}{\eta^T \eta} \sim (\tilde{X}^T \tilde{X})_{11} \sim \lambda \chi_p^2 \tag{28}$$

Hence by CLT,

$$\frac{\frac{\eta^T \tilde{X}^T \tilde{X} \eta}{\eta^T \eta} - \lambda p}{\lambda \sqrt{2p}} \xrightarrow{\mathcal{L}} N(0, 1) \tag{29}$$

$A_1 \asymp n \|\beta\|^2$. $A_3 \asymp n$. And $B_1 \asymp \frac{n}{p} \|\beta\|^2$. $B_3 \asymp n/p$.

$$\begin{aligned}
A_2 &= O_P(\|\beta\|/\sqrt{p}) \eta^T (\tilde{X}^T \tilde{X})^{1/2} \tilde{\epsilon} \\
&= O_P(\|\beta\|/\sqrt{p}) \sqrt{\eta^T (\tilde{X}^T \tilde{X}) \eta} \\
&= O_P(\|\beta\|/\sqrt{p}) O_P(\sqrt{np}) \\
&= O_P(\sqrt{n} \|\beta\|)
\end{aligned} \tag{30}$$

$$\begin{aligned}
B_2 &= O_P(\|\beta\|/\sqrt{p}) \eta^T (\tilde{X}^T \tilde{X})^{-1/2} \tilde{\epsilon} \\
&= O_P(\|\beta\|/\sqrt{p}) \sqrt{\eta^T (\tilde{X}^T \tilde{X})^{-1} \eta} \\
&= O_P(\|\beta\|/\sqrt{p}) O_P(\sqrt{n/p}) \\
&= O_P\left(\frac{\sqrt{n}}{p} \|\beta\|\right)
\end{aligned} \tag{31}$$

If $\|\beta\|^2 \rightarrow 0$, we have

$$\begin{aligned}
& \left| \frac{T - \lambda p}{\lambda\sqrt{2p}} - \frac{\frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} - \lambda p}{\lambda\sqrt{2p}} \right| \\
&= \frac{1}{\lambda\sqrt{2p}} \left| T - \frac{\tilde{\epsilon}^T \tilde{\epsilon}}{\tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \right| \\
&= \frac{1}{\lambda\sqrt{2p}} \left| \frac{A_1 + A_2 + A_3}{B_1 + B_2 + B_3} - \frac{A_3}{B_3} \right| \\
&= \frac{1}{\lambda\sqrt{2p}} \left| \frac{(A_1 + A_2)B_3 - (B_1 + B_2)A_3}{(B_1 + B_2 + B_3)B_3} \right| \\
&= \frac{O_P(1)}{\lambda\sqrt{2p}} \left| \frac{(O_P(n\|\beta\|^2) + O_P(\sqrt{n}\|\beta\|))O_P(\frac{n}{p}) - (O_P(\frac{n}{p}\|\beta\|^2) + O_P(\frac{\sqrt{n}}{p}\|\beta\|))O_P(n)}{n^2/p^2} \right| \\
&= O_P(\sqrt{p}\|\beta\|^2) + O_P(\frac{\sqrt{p}}{\sqrt{n}}\|\beta\|)
\end{aligned} \tag{32}$$

Hence if $\sqrt{p}\|\beta\|^2 \rightarrow 0$ and $p/n^2 \rightarrow \infty$, CLT holds.

If $\|\beta\|^2 \rightarrow \infty$, we have

$$\begin{aligned}
& \left| \frac{T - \lambda p}{\lambda\sqrt{2p}} - \frac{\frac{A_1}{B_1} - \lambda p}{\lambda\sqrt{2p}} \right| \\
&= \frac{1}{\lambda\sqrt{2p}} \left| \frac{A_1 + A_2 + A_3}{B_1 + B_2 + B_3} - \frac{A_1}{B_1} \right| \\
&= \frac{1}{\lambda\sqrt{2p}} \left| \frac{(A_2 + A_3)B_1 - (B_2 + B_3)A_1}{(B_1 + B_2 + B_3)B_1} \right| \\
&= \frac{O_P(1)}{\lambda\sqrt{2p}} \left| \frac{(O_P(\sqrt{n}\|\beta\|) + O_P(n))O_P(\frac{n}{p}\|\beta\|^2) - (O_P(\frac{\sqrt{n}}{p}\|\beta\|) + O_P(\frac{n}{p}))O_P(n\|\beta\|^2)}{\frac{n^2}{p^2}\|\beta\|^4} \right| \\
&= O_P(\frac{\sqrt{p}}{\sqrt{n}}\|\beta\|^{-1}) + O_P(\sqrt{p}\|\beta\|^{-2})
\end{aligned} \tag{33}$$

Hence if $\frac{n}{p}\|\beta\|^2 \rightarrow \infty$ and $p/n^2 \rightarrow \infty$, CLT holds.

3.2. circumstance 2

$$\begin{aligned}
T &= \frac{\beta^T \tilde{X} \tilde{X}^T \beta + 2\beta^T \tilde{X} \tilde{\epsilon} + \tilde{\epsilon}^T \tilde{\epsilon}}{\beta^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \beta + 2\beta^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon} + \tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \\
&= \frac{\beta^T P_2 P_2^T \tilde{X} \tilde{X}^T P_2 P_2^T \beta + 2\beta^T P_2 P_2^T \tilde{X} \tilde{\epsilon} + \tilde{\epsilon}^T \tilde{\epsilon}}{\beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T P_2 P_2^T \beta + 2\beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon} + \tilde{\epsilon}^T (\tilde{X}^T \tilde{X})^{-1} \tilde{\epsilon}} \\
&= \frac{A_1 + A_2 + A_3}{B_1 + B_2 + B_3}
\end{aligned} \tag{34}$$

If $\beta = 0$, then $T \sim (\tilde{X}^T \tilde{X})_{11 \cdot 2}$.

Let $\lambda_1 \geq \lambda_2 \dots \geq \lambda_p > 0$ be eigenvalues of Σ_X , where $\lambda_{r+1} = \dots = \lambda_p = \lambda$. Denote by $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$. Let $Z = (Z_1, \dots, Z_p)$ be a $n-1 \times p$ matrix with all elements independently distributed as $N(0, 1)$. Let $Z_{(1)}$ and $Z_{(2)}$ be the first 1 row and last $n-2$ rows of Z , that is

$$Z = \begin{pmatrix} Z_{(1)} \\ Z_{(2)} \end{pmatrix}.$$

Then

$$\begin{aligned}
\tilde{X}^T \tilde{X} &\sim Z \Lambda Z^T \\
&= \begin{pmatrix} Z_{(1)} \Lambda Z_{(1)}^T & Z_{(1)} \Lambda Z_{(2)}^T \\ Z_{(2)} \Lambda Z_{(1)}^T & Z_{(2)} \Lambda Z_{(2)}^T \end{pmatrix}.
\end{aligned} \tag{35}$$

Hence

$$\begin{aligned}
T &\sim Z_{(1)} \Lambda Z_{(1)}^T - Z_{(1)} \Lambda Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda Z_{(1)}^T \\
&= Z_{(1)} \Lambda^{1/2} (I_p - \Lambda^{1/2} Z_{(2)}^T (Z_{(2)} \Lambda Z_{(2)}^T)^{-1} Z_{(2)} \Lambda^{1/2}) \Lambda^{1/2} Z_{(1)}^T \\
&\leq Z_{(1)} \Lambda^{1/2} (I_p - \hat{V} \hat{V}^T) \Lambda^{1/2} Z_{(1)}^T.
\end{aligned} \tag{36}$$

We require $p = o(n^2)$. The principal space is $V = (e_1, \dots, e_r)$. Then

$$Z_{(1)} \Lambda^{1/2} (V V^T - \hat{V} \hat{V}^T) \Lambda^{1/2} Z_{(1)}^T = o(\sqrt{p}) \tag{37}$$

Note that

$$Z_{(1)} \Lambda^{1/2} (I - V V^T) \Lambda^{1/2} Z_{(1)}^T \sim \lambda \chi_{p-r}^2 \tag{38}$$

Hence $T \leq \lambda \chi_{p-r}^2 + o(\sqrt{p})$.

On the other hand, the eigenvalues of $\Lambda^{1/2}(I_p - \Lambda^{1/2}Z_{(2)}^T(Z_{(2)}\Lambda Z_{(2)}^T)^{-1}Z_{(2)}\Lambda^{1/2})\Lambda^{1/2}$ is no less than $I_p - \Lambda^{1/2}Z_{(2)}^T(Z_{(2)}\Lambda Z_{(2)}^T)^{-1}Z_{(2)}\Lambda^{1/2}$. Hence $T \geq \lambda\chi_{p-n+2}^2$.

Hence $A_3/B_3 \asymp p$ if $p/n \rightarrow \infty$.

Note that $P_2^T \tilde{X}$ is an $(p-r) \times (n-1)$ matrix with all elements independently distributed as $N(0, \lambda)$.

$$A_1 \asymp n\|P_2^T \beta\|^2, A_2 = O_P(\sqrt{n}\|P_2^T \beta\|), A_3 \asymp n.$$

$$B_3 \asymp n/p$$

As for B_1 ,

$$\begin{aligned} B_1 &\leq \beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T P_2 P_2^T \tilde{X})^{-1} \tilde{X}^T P_2 P_2^T \beta \\ &\asymp \frac{n-1}{p-r} \|P_2^T \beta\|^2 \end{aligned} \quad (39)$$

On the other hand, let $P_2^T \tilde{X} = U_2 D_2 V_2^T$ be the SVD of $P_2^T \tilde{X}$. Then

$$\begin{aligned} B_1 &= \beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T P_1 P_1^T \tilde{X} + \tilde{X}^T P_2 P_2^T \tilde{X})^{-1} \tilde{X}^T P_2 P_2^T \beta \\ &= \beta^T P_2 U_2 D_2 V_2^T (\tilde{X}^T P_1 P_1^T \tilde{X} + V_2 D_2^2 V_2^T)^{-1} V_2 D_2 U_2^T P_2^T \beta \\ &= \beta^T P_2 U_2 (V_2 D_2^{-1} \tilde{X}^T P_1 P_1^T \tilde{X} D_2^{-1} V_2^T + I_{n-1})^{-1} U_2^T P_2^T \beta \end{aligned} \quad (40)$$

Note that U_2 is independent of $(V_2, D_2, P_1^T \tilde{X})$, and

$$(V_2 D_2^{-1} \tilde{X}^T P_1 P_1^T \tilde{X} D_2^{-1} V_2^T + I_{n-1})^{-1} \geq I_{n-1} - U^* U^{*T} \quad (41)$$

where U^* is the first r eigenvectors of $V_2 D_2^{-1} \tilde{X}^T P_1 P_1^T \tilde{X} D_2^{-1} V_2^T$ and is independent of U_2 . Hence

$$\begin{aligned} B_1 &\geq \beta^T P_2 U_2 (I_{n-1} - U^* U^{*T}) U_2^T P_2^T \beta \\ &\asymp \frac{n-1-r}{p-r} \|P_2^T \beta\|^2 \end{aligned} \quad (42)$$

$$\begin{aligned} \text{Upper-Lower} &\leq \beta^T P_2 U_2 U^* U^{*T} U_2^T P_2^T \beta \\ &\asymp \frac{r}{p-r} \|P_2^T \beta\|^2 \end{aligned} \quad (43)$$

Hence

$$\begin{aligned} B_1 &= \beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T P_2 P_2^T \tilde{X})^{-1} \tilde{X}^T P_2 P_2^T \beta + O_p\left(\frac{r}{p-r} \|P_2^T \beta\|^2\right) \\ &= \beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T P_2 P_2^T \tilde{X})^{-1} \tilde{X}^T P_2 P_2^T \beta (1 + O_P(1/n)) \end{aligned} \quad (44)$$

$$\begin{aligned}
B_2 &= O_P(1) \sqrt{\beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T \tilde{X})^{-2} \tilde{X}^T P_2 P_2^T \beta} \\
&\leq \lambda_{\min}(\tilde{X}^T \tilde{X})^{-1/2} O_P(1) \sqrt{\beta^T P_2 P_2^T \tilde{X} (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T P_2 P_2^T \beta}
\end{aligned} \tag{45}$$

$$\lambda_{\min}(\tilde{X}^T \tilde{X}) \geq \lambda_{\min}(\tilde{X}^T P_2 P_2^T \tilde{X}) \asymp p - r \tag{46}$$

Hence $B_2 = O_P(\frac{\sqrt{n}}{p} \|P_2^T \beta\|)$.

Hence the similar law of large number and CLT holds.

3.3. Consistency of Test

β from normal distribution. Then consistency can be proved.

4. Simulation Results

References